Haplotype mapping in pharmacogenetics



- The identification of tagging SNPs in ADME genes
- 2. An application of haplotype tagging

A paradigm for indirect association studies (HapMap)



In this (unrealistic!) case the five haplotypes can be represented by just five tagging SNPs

- Genotype "control" individuals (LD data set)
- Assess LD and select tSNPs
- Type tSNPs in "phenotyped" individuals

Questions about ht mapping



- · How to select and test performance of tSNPs?
- · How many samples are required to select tSNPs?
- · How does performance depend on MAF?
- · How does performance depend on density?
- Is functional variation represented as well as non functional variation?
- How many more tSNPs are required to perform well in multiple populations?

Selection of Tags



- Select "tagging" SNPs that define G haplotypes that satisfy performance criteria in a regression model in which the G haplotypes are used as independent predictors to predict the allelic states of other "tagged" SNPs.
- $Y_i = x_{i1}b_1 + x_{i2}b_2 + ... + x_{iG}b_{iG}$

Refs: Weale ME et al. (2003) American Journal of Human Genetics 73: 551-565 Goldstein DB et al. (2003) Trends in Genetics 19: 615-622

Software implementation:

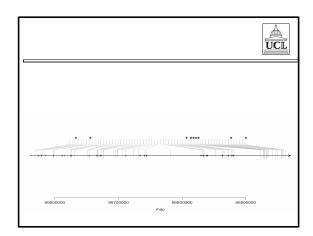
Weale ME & Goldstein DB (2003) TagIT software, http://popgen.biol.ucl.ac.uk

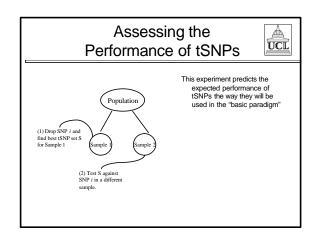
The data set ...

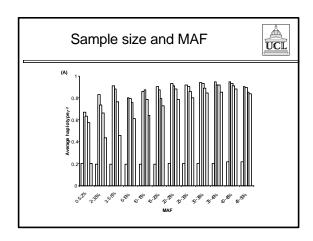


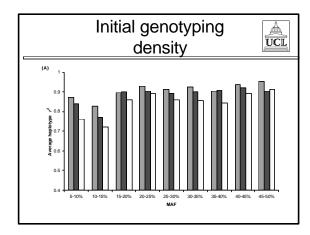
- 60 genes encoding most of the important DMEs
- Genotype approximately 1,000 SNPs in 64 CEPH and 64 Japanese individuals for a target density of one SNP / 2kb
 - -- GSK/UCL Collaboration --

				UCL
Family	Subfamily	Syronyms	Chrom	Total gene (b gsk
CYPIA	CYP1A complex (1.2)	cytochrome P450, family 1, subfamily A, polypectide 1/2	15022-024	27.592
CYPID	CYPIRI	cytochrome P450, family 1, subfamily II, polypeptide 1	2p21	21,904
CYP2A	CYP2A5	cytochrome P450, family 2, subfamily A, polypeptide 5	19q13.2	23,030
CYP28	CYP286	cytochrome P450, family 1, subfamily B, polypeptide 6	19013.2	30,250
CYP2C	CYP2C_complext (19,19)	cytochrome P450, family 2, subfamily C., polypeptide 19/19	10q24	212,629
	CYP2C_complex2 (9.8)	cytochrome P450, family 2, subfamily C , polypeptide B/D	90q24	140,295
CYP2D	CYP2D6 (Incl. D7P, D8P)	cytochrome P450, family 2, subfamily D. polypeptide GD7P1D8P	22q13.1	23,827
CYP2E	CYP2E1	cytochrome P450, family 2, subfamily E, polypeptide 1	10q24.3-qter	17,996
CYP2A	CYP3_complex (4.5.7, 5p1.5p2)	cytochrome P450, family 3, subfamily A, polypeptide 4/57/5p1/5p2	7g21.1	142,620
ALD	ALDH1A1	al dehyde dehydrog enas e 1 family, member A1	9q21.13	50,805
	ADH1C ALDH2	al cohol dehydrog enase 1C (class I), gamma polypeptide	4q21-q23	23,536
	ALDH2 NGC1	al dehyde dehydrog enase 2 family (mitochondrial) NADr Pirk dehydrog enase, quinone 1	12924.2	49,030 25,434
	DED (DEVD)	MALI(P)H denydrogenose, quinone 1	16q22.1 1622	25,494
NAT	NAT1	N-acetyltransferase 1 (andamine N-acetyltransferase)	9p23.1-p21.3	230,456 18,625
nem i	NAT1	N-acetyltransferase 1 (arytamine N-acetyltransferase) N-acetyltransferase 2 (arytamine N-acetyltransferase)	8p23.1-p21.3	18,625
OSTA	GGTA complex (1-2)	N-acetytransterase 2 (arytamine N-acetytransterase)	9902 9912.1	93 333
da i A	GSTA2	glutatione 5-transferane AV	0p12.1	16 949
	GSTA4	abilitions Stransferase A4	0012.1	23.837
SSTM	GGTM complex (1-5)	glutafrigne S-transferase M1,Q3,45	1012.2	99.717
DSTT	OSTT1	glutatrione S-tramferase theta 1	22011.22	14.020
	OSTT2	glutatrione 5-tramferane theta 2	22011.20	10,500
BBTZ	OSTZ1	glutsfrione transferose zeta 1 (malevlacetoscetate i somerase)	14924.3	18,207
JOT	UGT complex (1A/D.1A1.1A2.1A4.1A6.1A7.1	UDP glycosyltransferase 1 family, polypeptide A1/2 family, polypeptide 25	20.27	60,258
	ARSA	Antiufatme A	22g13.31-gter	12,249
	OPPO	Clucose-5-choschate dehydrogenase	Xq28	20,308
	ABCC1	ATP-binding cossette, sub-family C (CFTR/MRP), member 1	10p13:1	192,745
	ABCC2	ATP-tinding cossette, sub-family C (CFTR/MRP), member 2	10024	75,474
	ABCG1	ATP-binding consetts, sub-family G (WHITE), member 1	21q22.3	86,650
	ABC811	ATP-binding cossette, sub-family B (MDR/TAP), member 11	2124	114,462
	MVP	major voult protein	10p13;1-p11.2	28,153
	PON1 DON2	paraciicnase 1	7q21.3	32,137
		par acconase 2	7q21.2	22,526
	HNMT	Nistarrine N-methytransferase	2g22.1	50,411 22,425
	COMT	catechol-O-methyltransferase thickurine S-methyltransferase	22q11.21-q11.23 0s22.3	32,425
	NNMT	thiopurine 5-methytransferase ri optinamide N-methytransferase	0p22.3 11q23.1	29,309
	INMIT	ri cotinamide N-methytransferase frymidd ste synthetase	11q20.1 18p11.32	20,149
	PNMT	phenylethanolamine N-methytransferase	17q21-q22	11,428
TOTAL				2,145,803
MEAN				55.021



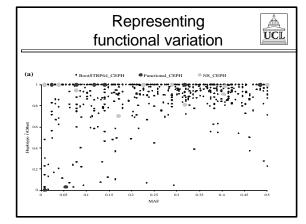






	Cosmopo	litan SNPs	UCL
• Europe	Э	196	
• Japan		179	

Cosmopolitan SNPs • Europe • Japan • Europe and Japan 226



Will tags hold in a different population



- Tags defined in CEPH
- Genotyped in a large sample from Aberdeen
- And tested against the panel of functional variants

Application in a new **UCL** population

Tagging the major human DMEs



- Density > 1 SNP / 5 kb results in detectable (but modest?) improvement in power
 Cosmopolitan tSNP sets require only modest increases in numbers over single population
- · Functional SNPs can be represented as well as non functional ones

Tagging the major human DMEs



- Density > 1 SNP / 5 kb results in detectable (but modest?) improvement in power
- Cosmopolitan tSNP sets require only modest increases in numbers over single population
- Functional SNPs can be represented as well as non functional ones
- · SNPs with low MAF are trouble!

Tagging the major human DMEs



 Approximately 200 SNPs are sufficient to represent the greater than 4,000 common polymorphisms in key genes regulating drug plasma levels

Phenytoin

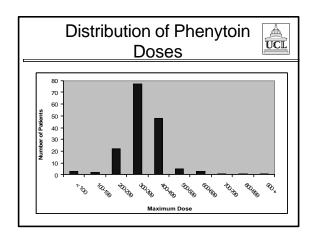


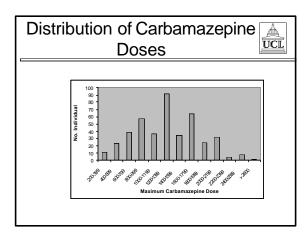
 A once widely used and effective anti epileptic drug that is now less often used (especially in Europe), in part because of adverse reactions (gum hyptertrophy, vision, headache)

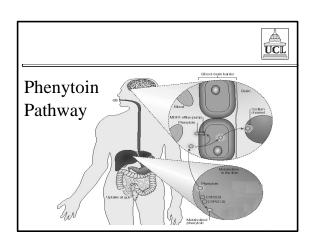
Carbamazepine



- Overall a safer drug than PHT, and more commonly used in the UK
- Identification of therapeutic dose can take months, and is complicated by tolerance





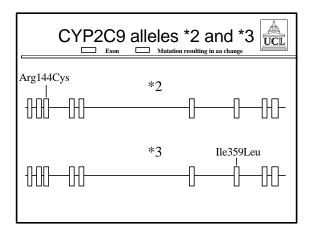


Strategy



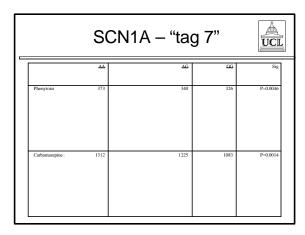
Phenytoin

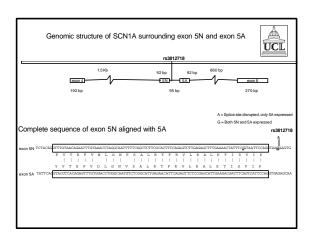
- Assess functional variants in CYP2C9
- · Assess putative functional variant in ABCB1
- Assess tagging SNPs in SCN1A Carbamazepine
- Assess putative functional variant in ABCB1
- Assess tagging SNPs in SCN1A
- Assess variation in PXR, regulator of CYP3A4



Representing variatiion in SCN1A				on
Gene Size	(Kb)	# SNPs	Density	# Tags
SCN1a	139	15	9.26667	4

_							
Results					UCI	A. 100	
	CYP2C9 Allele	Freq. (%)	Average Phenytoin Dose (mg)			Sig.	
		(,,,	wt/wt	wt/var	var/var		
	*2	11.9	357	339	317	0.38	
	*3	9.7	362	304	250	0.005	•
							•





Some good news (finally)...



Haplotype tagging can identify unknown functional variants

PHT and CBZ dosing



- PHT: Combining SCN1A and CYP2C9, the average maximum dose range between genetically sub-defined groups is 130 mg
- For CBZ, the average maximum dose range across SC1NA genotypes is 300 mg per day.

Clinical relevance?



- Rate at which AED dose can be safely increased
- Only prospective trials can determine whether PGx can improve clinical use will determine whether the clinical use of these medicines

Pgx is a simpler
complex trait



- Obvious candidate genes often carry PGx variants (drug specific)
- There may be direct clinical importance, although the demonstration will generally require prospective trials

Acknowledgments



- Tagging human drug metabolism
 - GlaxoSmithKline and University College London
- · Epilepsy genetics
 - Gianpierro Cavalleri, Chantal Depondt,
 Stephanie Schorge, Nicole Soranzo, Asra
 Siddiqui, Mike Weale,, Chantal Depondt, Ley
 Sander, Mari Wyn Burley, Richard Marguerie,
 Nicholas Wood, Sanjay Sisodiya